ED Assignment

Malik Jamie (JMXMAL003) and Amara Patel (PTLAMA005)

2025-10-21

Introduction

The aim of this experiment is to examine the effects of different intensities of heat and light on the growth of chilli plants. The goal is to determine whether light and heat affect chilli growth, which settings produce the highest quality chillies and by how much. Furthermore, whether these settings depend on variety as well as whether quality depends on variety will be explored. This is to make recommendations about which heat and light settings are optimal for small-scale farmers to implement, taking into account the cost of additional heat and light. The effects of various sources of potential variation - season, greenhouse and side of the greenhouse - will also be explored.

These goals will be tackled through exploratory data analysis, and an analysis of variance of this exploratory experiment.

Design and Randomisation

Design

The experiment takes place over 2 seasons – season 1 and season 2. There are two greenhouses available – greenhouse A and greenhouse B – with 8 north-facing and 8 south-facing plots each. The 32 experimental units were first blocked by the random factors, season and greenhouse, to give four blocks of 16 experimental units each. The treatment factors are side (north & south) and variety (Redhot & Furious), each with two levels which results in a 2^2 factorial treatment structure with 4 treatments: (n,F), (s,F), (n,R), (s,R).

Randomisation

Four 4x4 latin square designs were set up with the four levels of light (1, 2, 3, 4) as a row block and the four levels of heat (1, 2, 3, 4) as a column block. To conduct randomisation, a random latin square permutation of the four treatments, with each appearing once in every row and column, was found for each season-greenhouse combination using the design.lsd function from the agricolae package. The four treatments were randomly assigned to "a", "b", "c" and "d" for ease of understanding the latin square design layout.

Table 1: Randomly assigned treatment codes

treatment	code
nR	a
nF	b
sF	c
sR	d

Table 2: LSD for Greenhouse A, Season 1

	1	2	3	4
1	b	c	a	d
2	d	b	c	a
3	a	d	b	c
4	c	a	d	b

Table 3: LSD for Greenhouse B, Season 1

	1	2	3	4
1	b	d	c	a
2	a	c	d	b
3	d	a	b	c
4	c	b	a	d

Table 4: LSD for Greenhouse A, Season 2

	1	2	3	4
1	a	c	d	b
2	d	a	b	c

	1	2	3	4
3	b	d	c	a
4	c	b	a	d

Table 5: LSD for Greenhouse B, Season 2

	1	2	3	4
1	b	a	c	d
2	c	d	a	b
3	a	b	d	c
4	d	c	b	a

Hypotheses

The a-priori hypotheses that are of interest are:

1: Does the main effect of light have an effect on quality?

2: Does the main effect of heat have an effect on quality?

3: Does the interaction effect between light and heat have an effect on quality?

4: Does the interaction effect between light and variety have an effect on quality?

5: Does the interaction effect between heat and variety have an effect on quality?

6: Does the main effect of variety have an effect on quality?

Justification

The factorial treatment structure with side and variety – each with two levels crossed to give four treatment combinations – allows for the main effects of variety and side to be tested as well as the interaction effect between side and variety. This allows for the effect of variety on quality to be analysed.

The four latin squares, with light and heat as blocking factors with four levels each, allows for analysis of the variation due to light and heat in a balanced design, as there are four treatments that correspond to the four experimental units in each row or column with each treatment only appearing once in each column and row. This is also an efficient use of resources as 4 replicates of each side and variety combination are available per latin square. The latin structure also allows for the effects of light and heat on quality to be analysed thoroughly and easily.

Additionally, the factorial treatment structure in conjunction with the latin square design allows for the interactions between light and variety, and heat and variety to be analysed.

The replication across season-greenhouse blocks accounts for the variation due to season and the variation due to greenhouse, as well as ensuring sufficient power for the experiment as the 4 blocks result in $4 \times 4 = 16$ replicates per treatment.

Essentially, this design allows for efficient analysis of all the a-priori hypotheses of interest and enables many sources of variation to be accounted for.

Analysis and Results

Analysis

The model:

$$Y_{ijklmn} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_l + (\gamma\delta)_{kl} + \epsilon_m + \theta_n + (\epsilon\theta)_{mn} + (\delta\epsilon)_{km} + (\delta\theta)_{kn} + (\delta\epsilon\theta)_{lmn} + e_{ijklmn}$$
$$e_{ijklmn} \quad N(0, \sigma^2)$$

$$\begin{aligned}
Y_{ijklmn} &= \text{quality score for observation ijklmn} \\
\alpha_i &= \text{effect of } i^{th} \text{ season, } i = 1, 2 \\
\beta_j &= \text{effect of } j^{th} \text{ greenhouse, } j = A, B \\
\gamma_k &= \text{effect of } k^{th} \text{ side, } k = s, n \\
\delta_l &= \text{effect of } l^{th} \text{ variety, } l = F, R \\
\epsilon_m &= \text{effect of } m^{th} \text{ light level, } 1 \leq i \leq 4 \\
\theta_n &= \text{effect of } n^{th} \text{ heat level, } 1 \leq j \leq 4
\end{aligned}$$

An anova table was constructed based on this model with corner-point constraints for each variable as such:

$$\begin{aligned}
\alpha_1 &= 0 & \text{(Season 1 as reference)} \\
\beta_A &= 0 & \text{(Greenhouse A as reference)} \\
\gamma_n &= 0 & \text{(Side north as reference)} \\
\delta_F &= 0 & \text{(Variety Furious as reference)} \\
\epsilon_1 &= 0 & \text{(Light level 1 as reference)} \\
\theta_1 &= 0 & \text{(Heat level 1 as reference)}
\end{aligned}$$

Exploratory Data Analysis

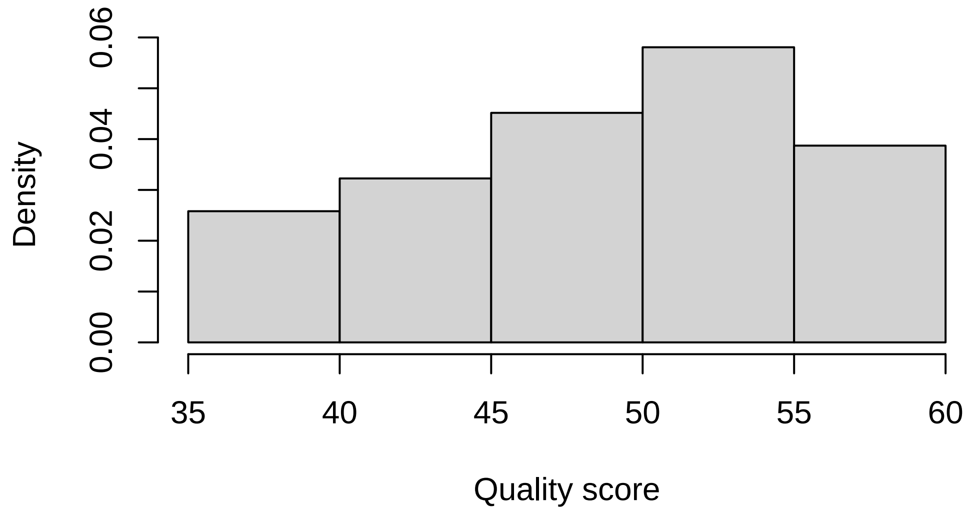

Figure 1: Histogram of quality score

The histogram of quality score suggests slightly left-skewed data.

Interaction Plots

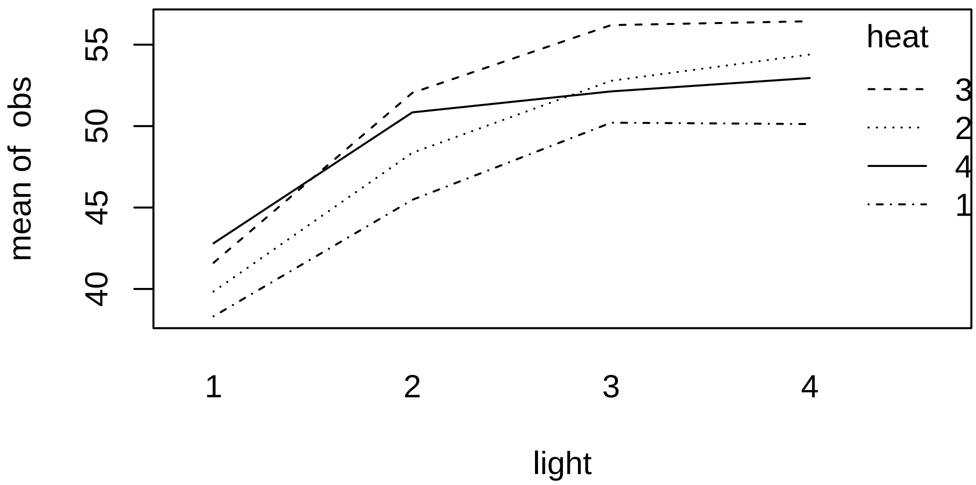

Figure 2: Interaction plot of light and heat

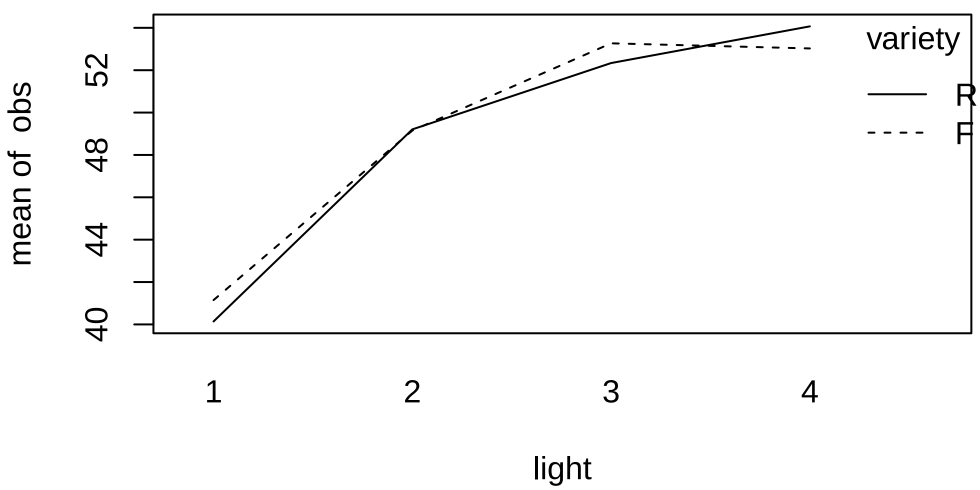

Figure 3: Interaction plot of light and variety

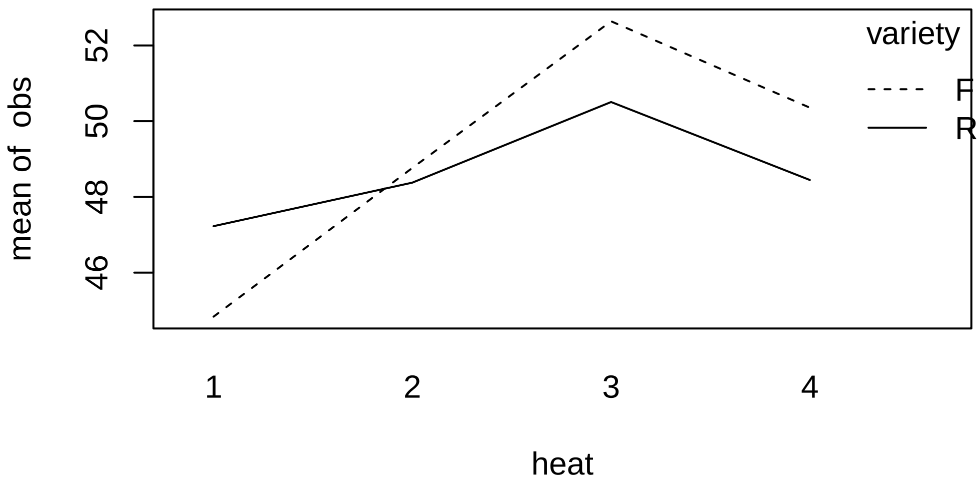

Figure 4: Interaction plot of heat and variety

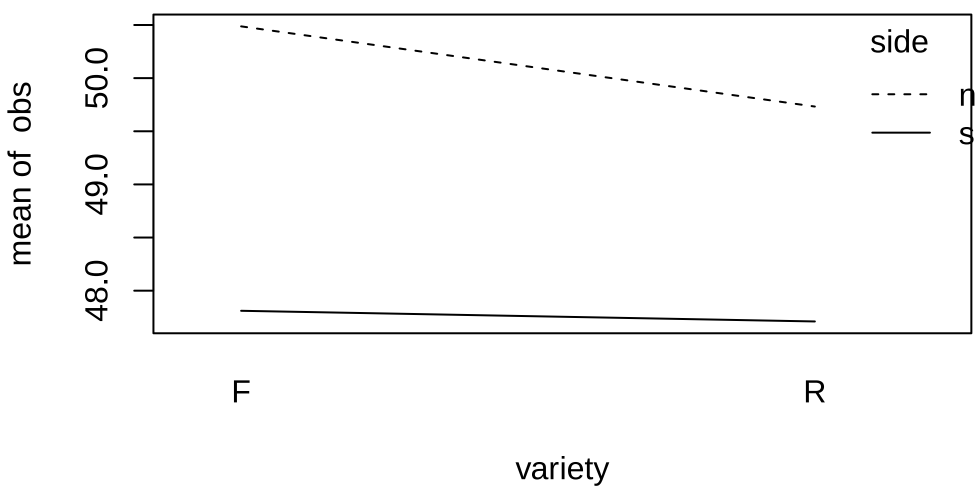

Figure 5: Interaction plot of variety and side

The interaction plot between light and heat suggests that there is minimal interaction between the four levels of each. The interaction plot between variety and side also suggests that there is minimal interaction between side and variety. On the other hand, the interaction plots between light and variety and heat and variety suggest some interaction effect.

Boxplots

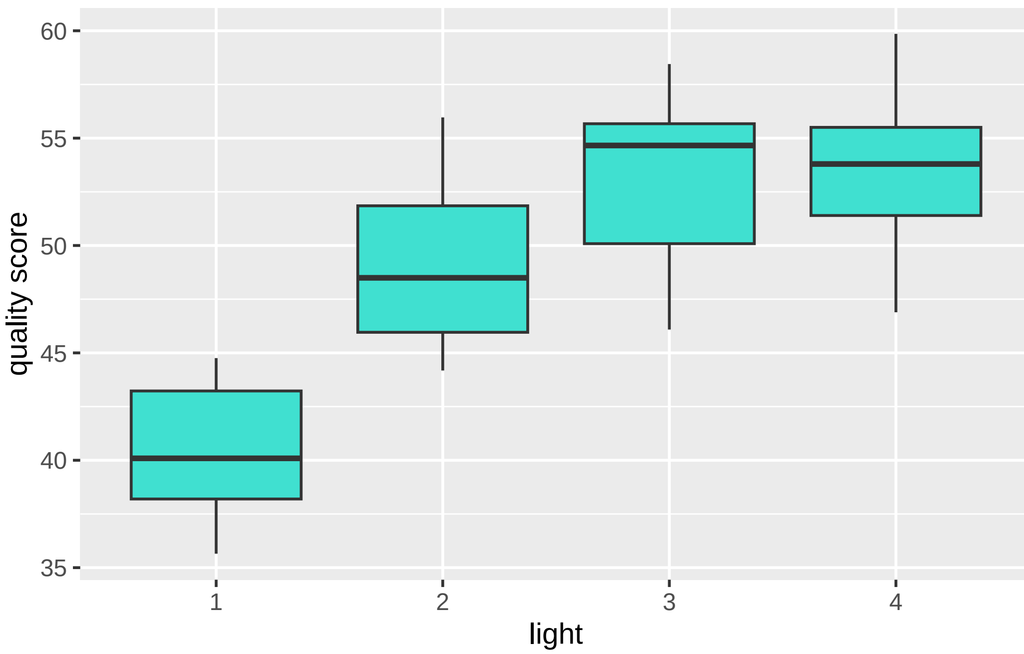

Figure 6: Boxplot of light levels

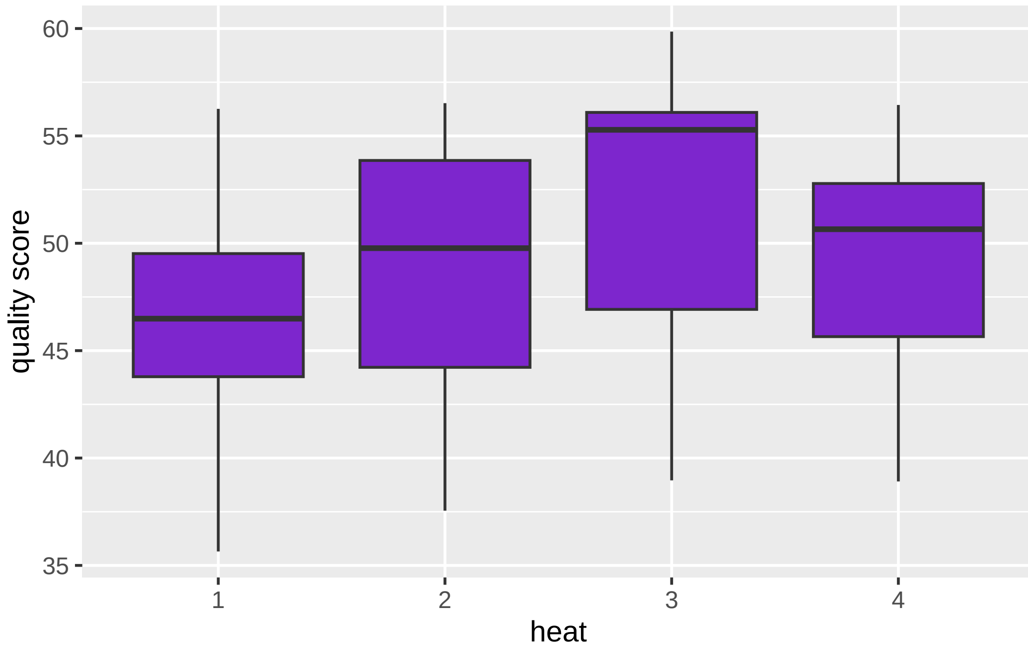

Figure 7: Boxplot of heat levels

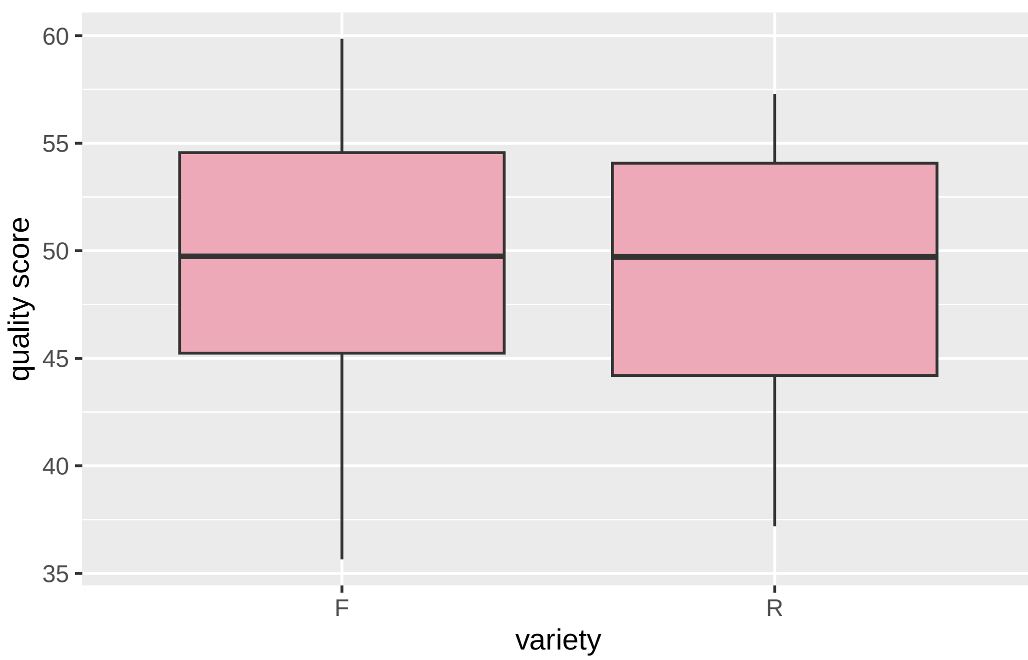

Figure 8: Boxplot of variety

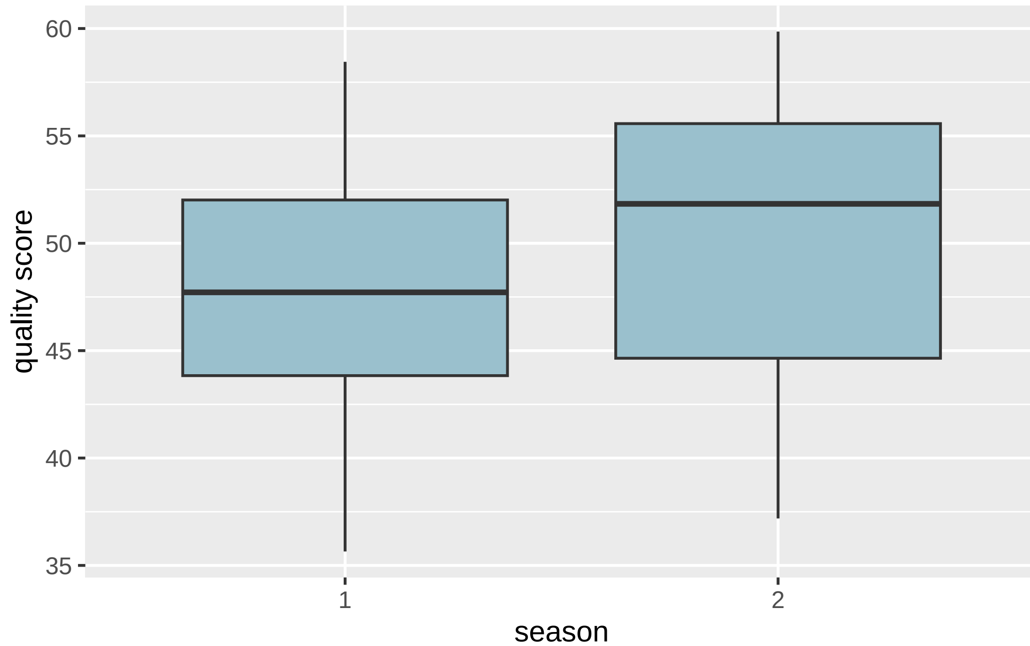

Figure 9: Boxplot of season

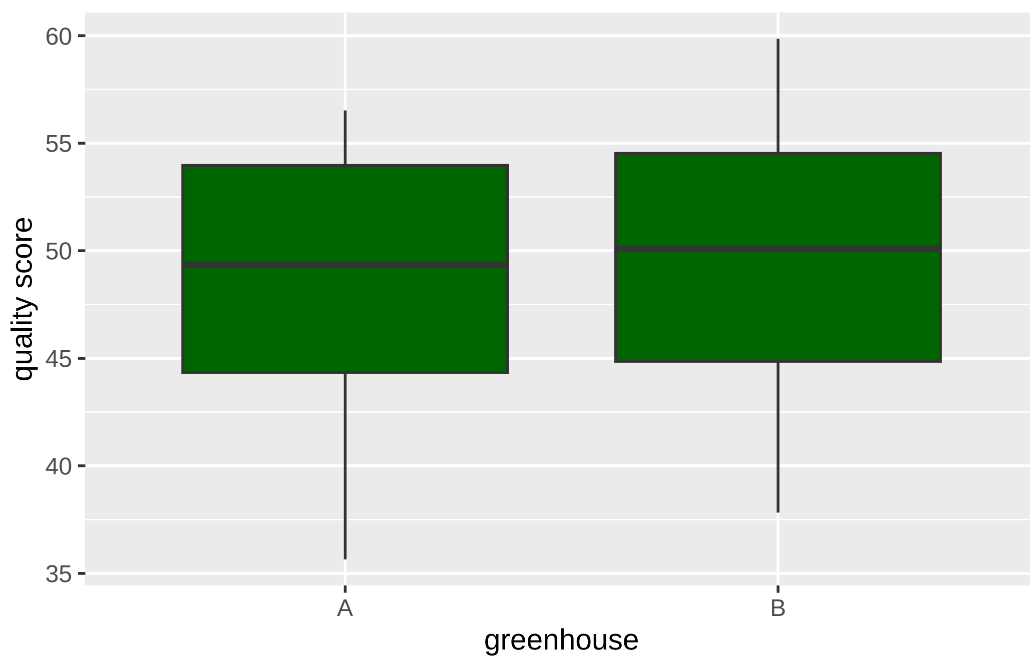

Figure 10: Boxplot of greenhouse

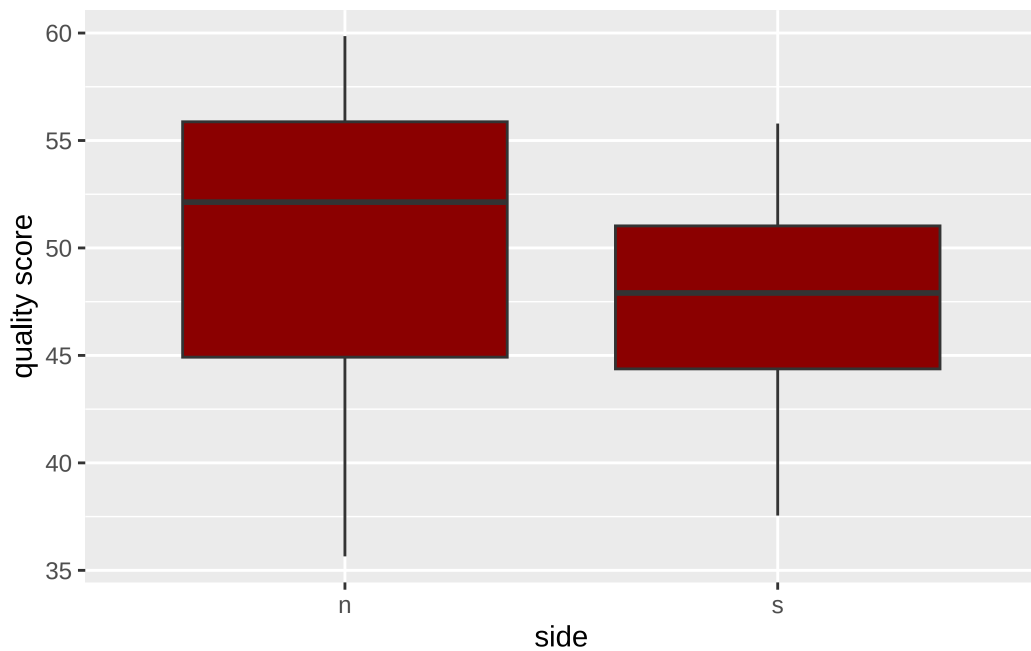

Figure 11: Boxplot of side

The boxplots show that, apart from the main hypotheses we are investigating: the effects of light, heat, variety and their interactions, there is a difference in median response from season 1 to season 2, as well as a difference in median response from the north side and south side of the two greenhouses.

Results

Table 6: ANOVA table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
season	1	122.5573291	122.5573291	56.719683	0.0000001
greenhouse	1	29.1017625	29.1017625	13.468332	0.0010989
variety	1	3.1280317	3.1280317	1.447657	0.2397461
side	1	89.7577718	89.7577718	41.540008	0.0000008
light	3	1679.3745913	559.7915304	259.072212	0.0000000
heat	3	261.7521490	87.2507163	40.379739	0.0000000
variety:side	1	0.0565276	0.0565276	0.026161	0.8727580
light:heat	9	26.0904991	2.8989443	1.341635	0.2643108
variety:light	3	38.3018161	12.7672720	5.908709	0.0032496

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
variety:heat	3	22.6533108	7.5511036	3.494660	0.0296904
variety:light:heat	9	28.2368503	3.1374278	1.452006	0.2178614
Residuals	26	56.1796253	2.1607548	NA	NA

Question 1: Does the main effect of light have an effect on quality?

Table 7: 95% Confidence Intervals of main effect of light levels in relation to light level 1

Light	Estimate	Lower bound	Upper bound
Level 2	7.157	3.2120	11.100
Level 3	11.330	7.7314	14.932
Level 4	14.040	10.9900	17.090

Our analysis provides strong evidence for the main effect of light having an effect on quality (p-val < 0.00001). The mean response with light level 2 was 7.157 higher than with light level 1, with a 95% confidence interval = (3.212, 11.102). The mean response with level 3 was 11.332 higher than with level 1 (7.731, 14.932). The mean response with level 4 was 14.043 higher than with level 1 (10.992, 17.094). Visually, the boxplots indicate that light level 3 actually had the highest median response, although level 4 had the highest mean response.

Question 2: Does the main effect of heat have an effect on quality?

Table 8: 95% Confidence Intervals of main effect of heat levels in relation to heat level 1

Heat	Estimate	Lower bound	Upper bound
Level 2	5.902	2.5690	9.235
Level 3	6.809	3.5185	10.099
Level 4	5.690	2.0900	9.291

Our analysis provides strong evidence for the main effect of heat having an effect on quality (p-val < 0.00001). The mean response with heat level 2 was 5.902 higher than with level 1 (2.569, 9.235). The mean response with heat level 3 was 6.809 higher than with level 1 (3.518, 10.099). The mean response with level 4 was 5.690 higher than with level 1 (2.090, 9.291). Visually, the boxplots indicate that heat level 3, as well as having the highest mean response, also had the highest median response.

Question 3: Does the interaction effect between light and heat have an effect on quality?

Our analysis does not provide sufficient evidence for an interaction between light and heat having an effect on quality (p-val = 0.240).

Question 4: Does the interaction effect between light and variety have an effect on quality?

Our analysis provides evidence for the interaction effect between variety and light (p-val < 0.005) having an effect on quality. Therefore, we proceeded with a contrast:

$$L_1 : \frac{1}{4}(\mu_{1R} + \mu_{2R} + \mu_{3R} + \mu_{4R}) - \frac{1}{4}(\mu_{1F} + \mu_{2F} + \mu_{3F} + \mu_{4F})$$

to test:

$$L_1 = 0$$
$$L_1 \neq 0$$

using a t-test.

Light	Variety	emmean
1	F	41.729
1	R	39.099
2	F	48.500
2	R	49.892
3	F	51.906
3	R	53.216
4	F	53.795
4	R	54.325

Table 10: Result of t-test for contrast L_1

Estimate	SE	Degrees of Freedom	t statistic	p-value
-8.506	0.4295	26	-19.8	0

Our t-test produces a statistically significant result (p-val < 0.0001). Therefore, there is strong evidence that the interaction effect between variety and light level has an effect on the quality of chillies.

Question 5: Does the interaction effect between heat and variety have an effect on quality?

Our analysis provides evidence for the interaction effect between variety and heat (p-val < 0.05) having an effect on quality. Therefore, we proceeded with a contrast:

$$L_2 : \frac{1}{4}(\mu_{1R} + \mu_{2R} + \mu_{3R} + \mu_{4R}) - \frac{1}{4}(\mu_{1F} + \mu_{2F} + \mu_{3F} + \mu_{4F})$$

to test:

$$L_2 = 0$$
$$L_2 \neq 0$$

using a t-test.

Table 11: Estimated marginal means for heat and variety combinations

Heat	Variety	emmean
1	F	45.153
1	R	45.914
2	F	48.218
2	R	49.155
3	F	52.250
3	R	50.077
4	F	49.830
4	R	50.920

Table 12: Result of t-test for contrast L_2

Estimate	SE	Degrees of Freedom	t statistic	p-value
-3.659346	0.4162093	26	-8.792081	0

Our t-test produces a statistically significant result (p-val < 0.0001). Therefore, we have strong evidence that the interaction effect between variety and heat has an effect on quality.

Question 6: Does the main effect of variety have an effect on quality?

Our analysis does not provide evidence for the main effect of variety having an effect on quality (p-val = 0.240). Visually, the boxplots indicate that the median responses for both varieties were very similar, when averaging across all other variables.

Further analysis

```
NOTE: Results may be misleading due to involvement in interactions

 light emmean    SE df lower.CL upper.CL
 1        40.5 0.402 26     39.7     41.4
 2        49.3 0.400 26     48.5     50.1
 3        52.7 0.449 26     51.7     53.6
 4        54.2 0.438 26     53.3     55.1

Results are averaged over the levels of: season, greenhouse, variety, side, heat
Confidence level used: 0.95
```

Conclusion and Recommendations

Conclusion

Your assignment is to design and analyse an experiment to help the two chilli farmers find the best settings for growing chillies and to help to understand how the requirements for the two varieties may differ.

The key aim of this assignment was to investigate the effects of different light and heat settings on the quality of chillies, and to determine which other factors influenced these results. We investigated the main effects of light, heat, variety, side, greenhouse and season on chilli growth, as well as various interaction terms between these different factors and treatments. Our choice of design was

Which light and heat settings produce the highest quality chillies?

From the ANOVA table, there was no evidence for the interaction term between light and heat being significant (p-val = 0.2643), nor the interaction term between light, heat and variety (p-val = 0.218). However, there is evidence for the interaction term between light and variety being significant (p-val = 0.0032), as well as the interaction term between heat and variety (p-val = 0.0297). Therefore, to determine the best light and heat settings, we calculated estimated means for heat and light for both varieties, averaging across all levels of other explanatory variables. Our results suggest that the settings that produced the highest quality chillies, on average, were different between the two varieties of chilli. For variety F, the light setting that produced the highest quality chillies on average, was level 4, with a mean response of 53.6 (52.6, 55.2). The optimal heat setting was level 3 with a mean response of 52.5 (51.3, 53.6). For variety R, the optimal light setting was level 4, with a mean response of 54.4 (53.1, 55.7). The optimal heat level 4, with a mean response of 51.1 (49.8, 52.4). Since the interactions between light, heat and variety were not significant, for each variety, the light and heat settings that would produce the highest quality chillies on average, were simply the settings that had the highest individual means. Therefore, for variety F, the optimal settings were light level 4 and heat level 3. For variety R, the optimal settings were light level 4 and heat level 4.